VICTORIA UNIVERSITY
MELBOURNE AUSTRALIA

Image preprocessing in classification and identification of diabetic eye diseases

This is the Published version of the following publication

Image Preprocessing in Classification and Identification of Diabetic Eye Diseases

Rubina Sarki[1] · Khandakar Ahmed[1] · Hua Wang[1] · Yanchun Zhang[1] · Jiangang Ma[2] · Kate Wang[3]

Abstract

Diabetic eye disease (DED) is a cluster of eye problem that affects diabetic patients. Identifying DED is a crucial activity in retinal fundus images because early diagnosis and treatment can eventually minimize the risk of visual impairment. The retinal fundus image plays a significant role in early DED classification and identification. An accurate diagnostic model's development using a retinal fundus image depends highly on image quality and quantity. This paper presents a methodical study on the significance of image processing for DED classification. The proposed automated classification framework for DED was achieved in several steps: image quality enhancement, image segmentation (region of interest), image augmentation (geometric transformation), and classification. The optimal results were obtained using traditional image processing methods with a new build convolution neural network (CNN) architecture. The new built CNN combined with the traditional image processing approach presented the best performance with accuracy for DED classification problems. The results of the experiments conducted showed adequate accuracy, specificity, and sensitivity.

Keywords Diabetic eye disease · Image processing · Convolution neural network

1 Introduction

Diabetic eye disease is the most common complication in diabetes, in which retinal fundus imaging is the most commonly adopted procedure because of its sensitivity in the diagnosis of DED [21]. The analysis of the severity and intensity of DED correlated with a patient having diabetes is typically attended by Ophthalmologists based on the lesion presents in retinal fundus images [50]. For instances, Fig. 1 presents the details on lesions that need to be identified from retinal images are; (i) Extra growth of blood vessels and damage or rupture in the tiny blood vessels in the retina (microaneurysms), often known as an early stage of diabetic retinopathy (DR); (ii) Built-up fluid causing swelling in the macular region or often form soft exedutes known as diabetic macular edema (DME). The common reason for blindness and vision loss, and (iii) Damage in the optic nerve and blood vessel rupture causes intraocular pressure that damages the optic nerve to inadequate, causing glaucoma (Gl). This eye condition is irreversible. While the number of diabetic patients is exponentially growing, there is also a rise in the number of retinal fundus images obtained by screening campaigns, which, in effect, causes a considerable labor-intensive and time-consuming complexity on the medical experts. This complexity has driven the development of an automated retinal lesions detection system.

Deep learning (DL) in machine learning (ML) has shown a significant impact on different science fields over the last few years [18, 22]. Such as the advancements in image and speech recognition, the ability to train artificial data that beat

✉ Jiangang Ma
  j.ma@federation.edu.au

  Rubina Sarki
  rubina.sarki@live.edu.vu.au

  Khandakar Ahmed
  khandakar.ahmed@vu.edu.au

  Hua Wang
  hua.wang@vu.edu.au

  Yanchun Zhang
  yanchun.zhang@vu.edu.au

  Kate Wang
  kate.wang@rmit.edu.au

[1] Victoria University, Ballarat Road, Melbourne, VIC 3011, Australia

[2] Federation University, Mount Helen, Australia

[3] RMIT University, Melbourne, Australia

Fig. 1 Early DED complication in retina; **a** Anatomical structure of the retina; **b** Microaneurysms- narrow buldges in blood vessels (diabetic retinopathy); **c** Soft exudates in macula (diabetic macular edema); (d) Optic nerve damage (glaucoma); and **e** cataract
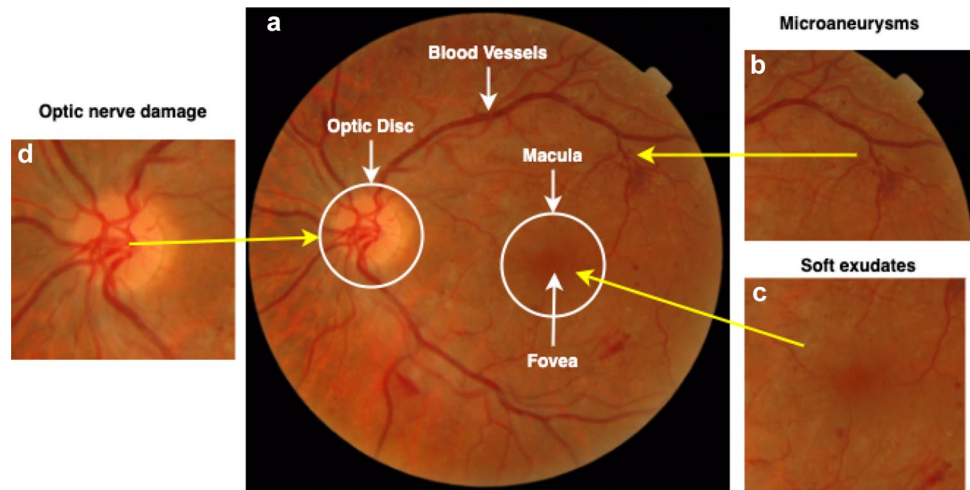
human players in games like ATARIS [30] and GO [42], and the development of new creative images using methods like Generative Adversarial Networks (GAN) [49], to learn the interaction between review texts and star ratings for prediction [12, 37] and as well as music [7]. Several of these activities were considered onerous to be accomplished by algorithms before the development of deep learning. Deep learning (DL) system plays a significant role in healthcare [11, 39, 48].

In deep learning, a pre-trained CNN network can be used to *transfer learning* from source task to target task with a limited number of images or minimise training time [28, 36]. The most popular *transfer learning* method is to fine-tune the pre-trained network. Regardless of the nature of the training model (pre-trained model or a new model), image data sets are typically preprocessed prior to training CNN architectures in various ways, such as image resizing, image quantity, image standardisation, and image enhancement. Improving the classification performance of the CNN model is limitless research, and the image quality in the data set has a significant impact on the overall performance of the architecture.

1.1 Motivation

Spectacular developments in the fields of artificial intelligence, computer vision and deep learning have led to remarkable results in image classification and vision tasks over the last several years, primarily through the use of DL [27]. However, in medical imaging (e.g., retinal fundus images), an early-stage identification of lesions and abnormalities is still the open issue reported in previous literature by Lam. et al. [27]. In their study they mentioned deep neural networks are struggling to learn enough in-depth features to identify aspects of mild disease; 93 percent of mild cases are wrongly classified as a healthy eye. Therefore this

research presents the system in which traditional image processing techniques and the state-of-the-art CNN combined to analyze early-DED disease. This paper articulates a research study using a small volume of the open-source retinal image database for in-depth learning evaluation between normal and mild DED classification.

1.2 Contributions

Therefore, in this research article, the main objective is to achieve the highest accuracy, sensitivity, and specificity than the existing deep learning models. The technique we used in this paper is the combination of traditional image processing methods for image enhancement and segmentation and then train in deep learning algorithms. We explore the significance of traditional image preprocessing in enhancing the early stage DED detection accuracy by using DL models. The advancement of this technology does not indicate the complete substitution of an ophthalmologist. Rather, it allows ophthalmologists more reliably diagnose DED. The contribution of this paper in the diagnosis of early DED can be classified into the following groups;

- Image enhancement: green channel extraction, contrast limited adaptive histogram equalization (CLAHE), and illumination correction was used to enhance the original image;
- Image segmentation: Region of Interest (ROI) such as blood vessels, macular region and optic nerve segmented from retinal fundus images;
- Pre-trained model: high-performance models were selected to classify the processed and segmented retinal fundus images;
- Build a new CNN model and training the model from scratch with processed and segmented retinal fundus images.

2 Related Works

Early detection of DED in retinal fundus images relies on a clinical technique to visualize a comprehensive set of features and localization within the image [4, 40]. Detection is challenging for diabetic patients with early DED stages because it depends on the existence of microaneurysms (bulges in blood vessels), fluid leakage from blood vessels, soft exudates formation, and damage of the optic nerve on retinal fundus images. The stages of diabetic eye disease are shown in Fig.1.

In the past, automated DED diagnostics have been explored to ease the burden on ophthalmologists and minimise diagnostic inconsistencies [31]. There are studies, which used lesion-based detection, an author like; Gharaibeh et al. [16] presented a new approach to detect microaneurysms in retinal fundus images. Their work includes preprocessing methods like blood vessel segmentation, fovea localization, and elimination. Then they used a combination of neural networks and fuzzy logical models for feature extraction and classification. Their study addressed the binary classification of diabetic retinopathy into two categories (microaneurysms and non-microaneurysms).

Moreover, a range of several other features than microaneurysms are appropriate for the diagnosis of DED. Similarly, Kaur et al. [25] proposed region-based segmentation and detection of the lesion and then classified using pixel-based classification and determine the severity level of the retinal disease. Karegowda et al. [24] detected exudates in diabetic retinopathy using decision tree and GA-CFS techniques as input to backpropagation neural network. They classified the normal eye and eye with exudate. The results obtained were insufficient to provide reliable classification accuracy and do not result in efficient noise removal. Sopharak et al. [47] presented a fuzzy c-means and clustering-based exudate identification method. Their work mostly relies on the identification of optic disc and the elimination of blood vessels. According to the results obtained, the exudates are identified without their characteristics. Jenuja et al. [23] present a method based on the optic disc and optic cup segmentation. The proposed method uses dual neural networks that operate in combination with the optical cup and disc parts. The aim of this proposed method is to efficiently segment the optic cup and disc of a retinal fundus image. The results of the classification of various stages of glaucoma are not given. Earlier Gulshan et al. [17] and Gargeya et al [15] presented CNN for DR detection using fundus images. They achieved specificity and sensitivity in the range of 90 percent for (normal/mild to moderate/severe) binary classification in private wider data sets comprising 80,000 to 120,000 fundus images.

There are many traditional strategies for DED diagnosis and classification [8, 19, 51]. Most techniques use neural networks, mathematical morphology, region of interest techniques, pattern recognition, clustering of fuzzy C-means, and Gabor filter techniques. For example, Chaudhuri et al. [8] uses 2D matched filters, to detect the blood vessels present in the retina. Vallabha et al. [51] uses Gabor filter bank outputs to identify the mild, moderate, and extreme stages of retinopathy, the automated detection and classification of abnormalities present in the vascular network are carried out.

There are numerous methods suggested for optic disc detection. Principal component analysis (PCA) is one of the methods by which clustering of brighter pixels derives the candidate regions for the optical disc. Noronha et al. [32] used Hough Transform for Optic Disc detection. In the detection of exudates, a neural network-based approach is used by Gardner et al. [14], For the detection of exudates, a fuzzy C-means clustering method employed by bezdek et al. [5] and a computational intelligence-based approach are used by Osareh et al. [33]. The automatic classification of normal, mild, moderate, severe, and proliferative diabetic retinopathy is carried out by measuring the areas of several characteristics, such as haemorrhages, microaneurysms, exudates, and blood vessels classified by the support vector machine [2].

However, the accuracy metrics for the diagnosis of four categories of DR, i.e., no DR, mild, moderate, and severe, are significantly based on disease-grade selection ratios. Although the no DR and severe stages are likely to achieve high sensitivity, the mild and moderate recall levels are often deficient. Research studies using publicly available datasets reveal difficulties in detecting early-stage DEDs.

3 Methods

The research's overarching objective is to improve the performance of early detection of DED from fundus images through the empirical assessment of image preprocessing and classification improvement techniques. The related objectives can be described as follows:

- Implementing traditional image processing techniques such as (i) Image Enhancement, (ii) Image Augmentation, (iii) Image Segmentation;
- Implementing various hyperparameters and evaluate their effect on CNN models' performance;
- Evaluate the accuracy obtained by pre-trained CNN models: *ResNet50*, *VGG-16*, and *Xception* with original and preprocessed fundus images;
- Developing a new CNN model to train preprocessed fundus images for classification accuracy improvement;
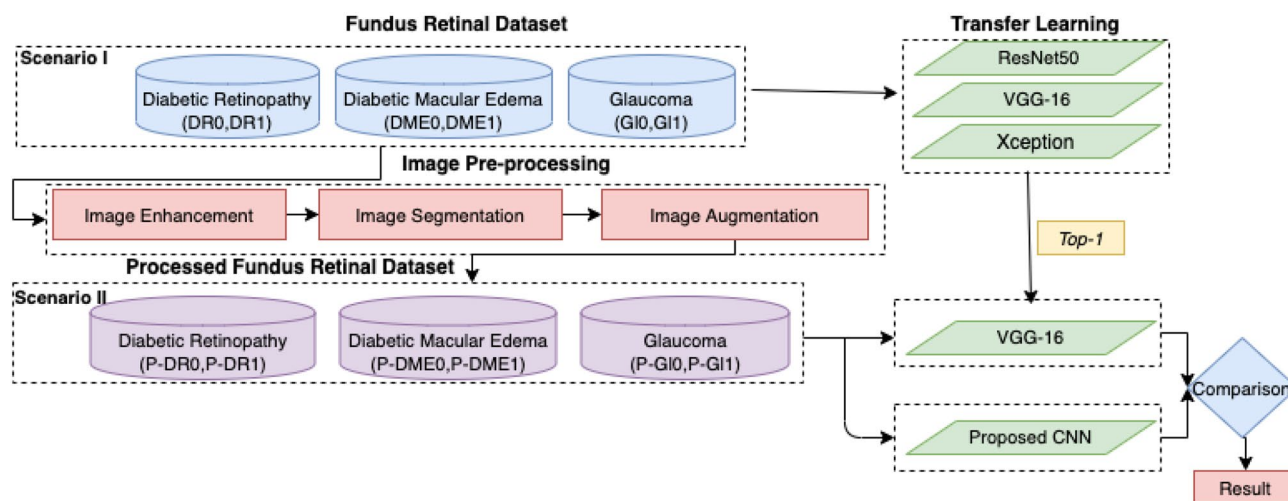
Fig. 2 The high-level process pipeline

- Evaluate the results of pretrained and new CNN model by performance metrics.

To high-level process, a pipeline is shown in Fig. 2 which demonstrates the importance of this workflow. In which, we experimented with the raw retinal fundus dataset employing three pre-trained models such as, (ResNet50, VGG-16, Xception) and obtained the highest performing model. We used traditional image processing algorithms to our raw fundus images and then trained this dataset with best performing model from the previous experiment. We also trained preprocessed images with CNN architecture from scratch. Finally, we compared the results to check preprocessed images improved the performance accuracy of the models or not.

3.1 Data Collection

Data was collected from publicly accessible sources, i.e., Messidor, Messidor-2, DRISHTI-GS, and Retinal Dataset from GitHub. This section explains the data sets used in these articles. The labeling of each image is generated by the ophthalmologist. Depending on the number of hemorrhages, microaneurysms, and the presence of neovascularisation, each image is classified as one of three lesion grades. *Messidor Dataset* has been formed to promote computer-assisted DED studies. Messidor database collected 1200 retinal fundus images of the posterior pole from three departments of ophthalmology using a 3CCD color video camera placed on a Topcon TRC NW6 non- retinograph with a 45° field of view (FOV). The medical experts offered two diagnostics for each image: *Retinopathy grade* and *Macular edema risk*. *Messidor-2 Dataset* is a publicly accessible dataset used by other individuals to evaluate DED algorithms' performance.

Messidor-2 comprises 1,748 colour retina images of 874 subjects. Messidor-2 varies from the actual Messidor dataset of 1200 images and ensures that it has two images for each subject, one for each eye. Using the previously published ICDR and DME gradings, Messidor-2 provided four disease rates for each subject. *DRISHTI-GS Dataset* [44] There are 101 retinal images in the Drishti-GS1 dataset with 31 normal images and 70 glaucoma lesion images. Due to the limited images obtained from DRISHTI-GS, we considered the glaucoma dataset from GitHub [1] *Retina Dataset* which contained 100 retinal images indicating glaucoma lesions. Therefore, improvement in the imbalance dataset can cause the improvement of the predictive model [52]. Thus we performed undersampling of the dataset in this study and selected 100 images from each class to perform our experiment.

3.2 Image Pre-processing

The preprocessing step is used to eliminate noise/variation in the retinal fundus image and improve the quality and contrast of the image. Apart from contrast enhancement and noise reduction, the preprocessing step can be used for image normalization and non-uniform intensity correction to eliminate artifacts and increase the accuracy of the process steps. Furthermore, DED features are localized, extracted, and segmented from fundus images for further classification in pre-trained models. The preprocessing techniques utilized in this article are briefly discussed in this section.

3.2.1 Image Enhancement

To enhance the original images' appearance and information value before processing, we used image enhancing techniques, popular techniques such as contrast enhancement, illumination correction. *Contrast enhancement*: Contrast limited adaptive histogram equalization (CLAHE) [56] is utilize to improve the visibility of images. CLAHE is an adapted part of the Adaptive Histogram Equalization (AHE) process. In this method, the enhancing function is introduced to all neighborhood pixels, and the transformation function is derived. This is distinct from AHE for its limited contrast. In CLAHE, the contrast of the image is improved by implementing Contrast limited histogram equalization (CLHE) to small data areas called tiles rather than the entire image. The resulting adjacent tiles are then perfectly stitched back utilising bilinear interpolation. CLAHE applied to greyscale retinal images. The 'cliplimit' function is applied to limit noise in an image. Create gray level mapping and clip the histogram. In the contextual area, pixel numbers are divided equally at each grey level so that the average number of pixels is grey as follows:

$$n_{avg} = \frac{n_{CR-x_p} * n_{CR-y_p}}{n_{gray}} \tag{1}$$

Where, n_{avg} = average number of pixels, $n_{g}ray$ = number of gray level in contextual region

n_{CR-x_p} = number of pixels in x direction of contextual region

n_{CR-y_p} = number of pixels in y direction of contextual region After that calculate the actual cliplimit.

$$n_{CL} = n_{CLIP} * n_{avg} \tag{2}$$

CLAHE [56] is an useful technique in biomedical image processing because it is very effective at making the normally important salient sections more accessible.

Illumination correction This preprocessing method aims to reduce the scenario effect caused by uneven illumination of retinal images [13]. Every pixel intensity is calculated using the following equation:

$$p' = p + \mu_D - \mu_L \tag{3}$$

Where p, p'is the initial and the latest pixel size values, respectively, μD is the desired average intensity, and μL is the local average intensity [54]. Microaneurysms forming at the surface of the retina are enhanced by this method.

3.2.2 Image Segmentation

To build an effective deep learning-based classification system for detecting mild DED, we need to consider the importance of the architecture of the network as well as the importance of input data. To obtain an efficient results, input images plays significant role. In retinal fundus images, variability such as the number of images, luminosity, contrast, and anatomical features determines the forthcoming result of the automatic disease detection algorithm. Therefore, features segmentation enhances the value of the images for classification and contribute for better accuracy. The process and the necessary theory is explained in the following sections.

Blood Vessels Extraction For detecting early stages of DR, the blood vessels are one of the most significant anatomical feature in retinal images. Thus retinal blood vessels segmentation is performed with following steps: (1) image enhancement; (2) Tyler Coye algorithm [10] and (3) morphological operation for further improvement in results.

We performed image enhancement techniques as mention above, green channel of the RGB color space presents better contrast between vessels network and background. The variation of contrast and luminosity in a background of a fundus image, can be estimate using method introduced by Zuiderveld [56] and Youssif et al. [54]. After, contrast and luminosity adjustment, ISODATA used in Tyler Coye algorithm is used for extracting threshold level. After tyler coye algorithm, morphological operation (erosion and dilation) used for further enhancement. Using these two essential fundamental operations we reduce noise or remove of gaps in the background and foreground. Erosion is a procedure used to eliminate or spike the edge of the area, which is represented in the following equation.

$$A \ominus B = \{p | B_p \subseteq A\} \tag{4}$$

Dilation is a procedure employed to broaden the rim of the background or foreground image configuration. This procedure is widely used to fill a gap, that can be defined in the following equation.

$$A \oplus B = \{x | B_x \cap X \neq 0\} \tag{5}$$

Closing is to perform the dilation, followed by erosion, to create a relation between each pixel of the image in order to bring it closer to one another. This procedure can be defined in the following equation,

$$A \cdot B = (A \oplus B) \ominus B \tag{6}$$

Fig. 3 The flow chart
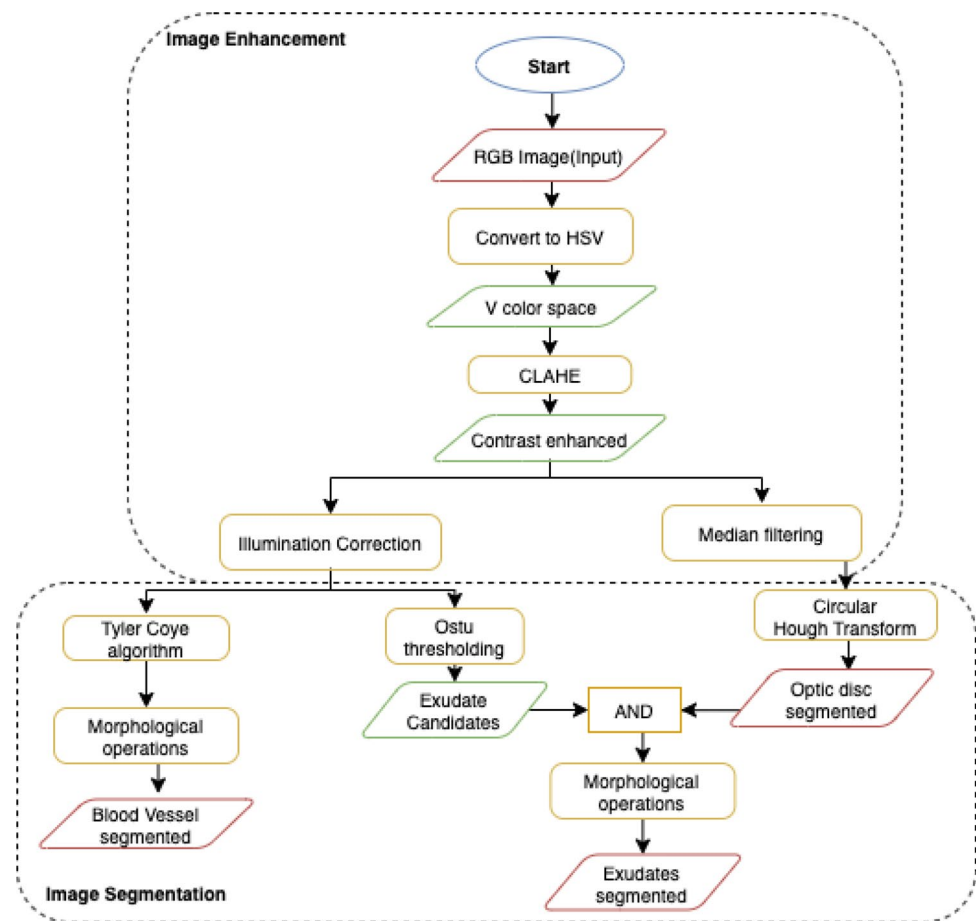
Where, $\oplus$ denote the dilation; $\ominus$ denote the erosion; $A =$ Structuring element and $B =$ the erosion of the dilation of that set. However, several gaps remain in Tyler Coye algorithm. This morphological process is to fill these small gaps in order to cover some of the required regions of the blood vessels.

Optic Disc Detection and Extraction: Glaucoma occurs when optic nerve is damaged. So, the segmentation of the optic disc (OD) helps to view the clearer anatomical changes in optic nerve. Figure 1 shows an fundus image of an eye from our collected data set with anatomical parts showing OD. To segment the OD we applied following steps: (i) image enhancement, (ii) Circular Hough Transform (CHT) to detect circular object, (iii) median filter to reduce noise, and (iv) optic disc segmentation using the threshold values. Image processing attempts to improve and increase the quality of the retinal fundus image in order to enable the identification of clinical features for DED. Flowchart of the image processing and image segmentation approach is depicted in Fig. 3. CLAHE can not be employed in the entire image, but only on a specific area 'tile' of the image.

Image enhancement calculation is adjusted on the basis of the user-specific maximum contrast-rate level by setting

its rate to $l, 0 \leq l \leq l$ [46]. Further contrast enhancement is performed in those images which has low contrast estimated by

$$\phi(i,j) = \left(\frac{\mu(i,j) - \Delta}{\delta - \Delta} \right)(\Gamma - 1) \tag{7}$$

Where, $\phi(i,j)$ and $\mu(i,j)$ are pixels after transformed and pixels before transformed in (i,j) coordinate, respectively; Δ is maximum pixel value; δ is minimum pixel value of input image and Γ is maximum value of gray scale.

The median filtering has a strong noise reduction efficiency and it is very common in image processing for noise removal. Mean filtering replaces the pixel value in the middle of the sliding window with the median value of the pixels in the window. Mathematically median filtering is represented in,

$$f(x,y) = median_{(s,t) \in S_{xy}} \{g(s,t)\} \tag{8}$$

Segmentation is a pixel classification method for extracting objects or segmenting regions with a similar attributes from the background [41]. Therefore, we used the Circular Hough Transform (i.e. CHT) method for optical disc detection. The CHT method is often used to identified circular shape in

an image. The key benefit of the CHT approach is that it is sensitive of differences in feature specification descriptions as well as largely unaffected by image noise. The CHT is provided by given equation:

$$(x - a)^2 + (y - b)^2 = c^2 \tag{9}$$

The procedure to detect circles involves the following steps:(i) we obtain a binary edge map of the image; (ii) values for a and b are set; (iii) obtain the value of c radius that satisfies Eq. 7; (iii) adjust the accumulator corresponding to (a,b,c); (iv) change the values for a and b within the scope of interest and return to Phase (iii).

Exudate localization and Detection: Exudate in two-dimensional retinal images acquired via a digital fundus camera, usually appear as bright area with varying scale, brightness, position and form. Precise exudate segmentation is a difficult activity given the large variety of scale, intensity, contrast and shape. It comprises of three major processing stages: (1) image enhancement; (2) optic disc detection and removal; (3) blood vessel removal; and (4) exudate extraction. When exudate are acquired from the mild dataset, the classification of the DME can be performed according to the grading criteria mention in the messidor dataset. Early DME can be diagnose early by detecting the presence of exudates in fundus images. Figure 1 shows the exudates formation in the macular region. After *optic disc detection and removal* performed, Otsu thresholding is applied to obtain candidate areas of exudates. Threshold value T relying on the input image is estimated by Ostu method, automatically. First, intensity value i of histogram is calculated using Eq. (11),

$$p(i) = \frac{n_i}{N}, p(i) \geq 0, \sum_1^{256} p(i) = 1 \tag{10}$$

The number of pixel images N and the number of pixels n_i with I intensity. Subject weight and background are described in (12) and (13).

$$w_1(t) = \sum_{i=1}^{t} p(i) \tag{11}$$

$$w_2(t) = \sum_{i=t+1}^{L} p(i) = 1 - w_1(t) \tag{12}$$

Here, the number of the gray level is L. The mean of the object and the background is then determined using Eqs. (14) and (15)

$$m_1(t) = \sum_{i=1}^{t} i.p(i)/w_1(t) \tag{13}$$

$$m_2(t) = \sum_{i=1}^{t} i.p(i)/w_2(t) \tag{14}$$

Hence, Variance is estimated by Eqs. (16, 17), while total of variance is expressed in Eq. (18) as follows.

$$\sigma_1^2(t) = \sum_{i=1}^{t} (1 - m_1)^2 . \frac{p(i)}{w_1(t)} \tag{15}$$

$$\sigma_2^2(t) = \sum_{i=t+1}^{t} (1 - m_2)^2 . \frac{p(i)}{w_2(t)} \tag{16}$$

$$\sigma^2(t) = \sigma_w^2(t) + \sigma_B^2(t) \tag{17}$$

Here, $\sigma 2w$ is called as within-class variance (WVC) that is expressed in Eq. (19), while $\sigma 2B$ called between-class variance (BVC) that is expressed in Eq. (20). WVC is the amount of individually class variance that has been weighted with probability of each class. Average total is calculated using Eq. (21). Threshold value can be obtained from minimisation of WVC or maximisation of BVC, but BVC has less computation time.

$$sigma_w^2(t) = w_1(t).\sigma_1(t)^2 + w_2(t).\sigma_2(t)^2 \tag{18}$$

$$\sigma_B^2(t) = w_1 . [m_1(t) - m_T]^2 + w_2 . [m_2(t) - m_T]^2 \tag{19}$$

$$m_T = \sum_{i=1}^{N} i.p(i) \tag{20}$$

Morphological is a set of discrete coordinates that related to pixel object of image that involves logical operation, such as "or" and "and". Opening operation aims to refine object contour and repair object contour with eliminated pixel area that smaller than structure element. Opening operation is expressed in Eq. (22).

$$AoB = (A\Theta B) \oplus B \tag{21}$$

3.3 Image Augmentation

Deep learning models perform well with high-volume training data [26, 35]. Therefore, data augmentation comprises a collection of techniques that improve the quantity of training data without actively acquiring new data. Thus, image augmentation algorithms addressed in this paper include geometric transformations such as flipping, rotation, mirror, and cropping. We used Keras *ImageDataGenerator* class for real-time image augmentation, which ensures that the selected model will obtain variations of the images at every

Table 1 Three CNN models pre-trained on ImageNet and their characteristics. *Source*: https://keras.io/applications)

Model	Size	Top-1 accuracy*	Top-5 accuracy*	Parameters	Depth**	Reference
Xception	88 MB	0.790	0.945	22,910,480	126	Chollet [9]
VGG16	528 MB	0.713	0.901	138,357,544	23	Simonyan and Zisserman [43]
DenseNet21	33 MB	0.750	0.923	8,062,504	–	Huang et al. [20]

*The top-1 and top-5 accuracy refers to the model's performance on the ImageNet validation dataset.

**Depth refers to the topological depth of the network. This includes activation layers, batch normalisation layers etc
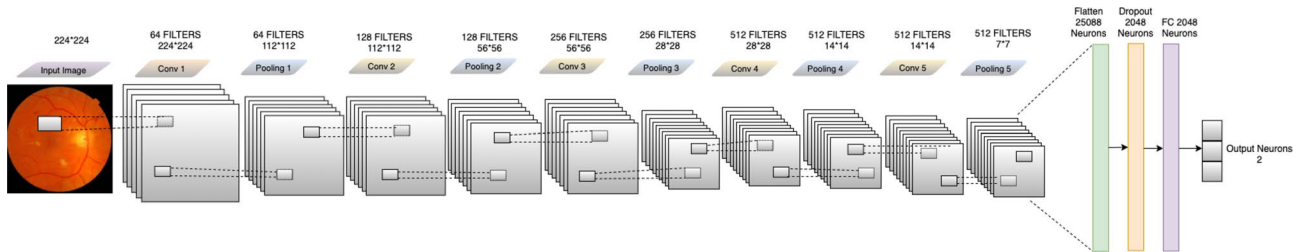

Fig. 4 The proposed CNN model

epoch. The advantage of using *ImageDataGenerator* class in our work is transformed images will not add to the range of original images, which avoid overfitting the selected model.

3.4 Transfer Learning

In this research, we are using CNNs-based transfer learning to implement the DED retinal fundus image classification. To accomplish the absolute best classification outcomes, we explore pre-trained CNN model transfer learning techniques. The precise details of the pre-trained models will be presented in this section.

According to Pan et al. [34] transfer learning is defined as; $D = \Phi, P(X)$ with $X = x_1, x_2, ..., x_n \, \epsilon \, \Phi$, where, D is domain, Φ is refer to a feature space and $P(X)$ marginal probability distribution. Given, $T = Y, F(*)$ where, T is given task, Y is refer to a label space and $F(*)$ is an objective predictive function that is learned from the feature vector and label pairs. Specifically, given a source domain D_s with learning task T_s and a target domain D_t with learning task T_t, then transfer learning is the process of improving the learning of the target predictive function $F_t(*)$ in D_t based on the knowledge learned from source domain D_s and learning task T_s, where $D_t \neq D_t$, or $T_s \neq T_t$. It should be noted that the single source domain described above can be expanded across multiple source domains.

The concept behind the transfer learning for the classification of images is that if a network is typically trained on a broad scale and enough data set (e.g., ImageNet), it can effectively train in the particular target task, which has fewer

labeled examples than the pre-training dataset. One can benefit from these learned feature maps without training a large model from scratch on a large dataset.

In this paper, we will customize pre-trained models in two ways: *(1) Feature Extraction*, features learned from the source task to extract useful features from the target task. We added a new classifier, which can be trained from scratch to the top of the pre-trained network to modify the features maps initially learned for the sample. *(2) Fine-Tuning*, unfreeze some of the last layers of the frozen base network and collectively train the last layers of the base network and the newly added classifier layers. This helps to "fine-tune" the higher-order character representations in the base network to make them more appropriate to the target task. We fine-tune three pre-trained CNN networks(Xception, VGG-16, and DenseNet21) to implement DED image classification. Three CNN pre-trained network on ImageNet and their characteristics are described in Table 1.

3.5 Proposed CNN Model

CNN's are the most popular deep learning algorithms which train the medical images for the classification of medical image abnormalities [29]. The explanation for this is while analyzing input images, CNN preserves distinctive features. Spatial relationships, such as where the blood vessels start rupturing or how yellow fluid starts accumulating near the macular region, are of primary importance in retinal images, as we discussed above. The framework of the process is shown in Fig. 4 and Table 2 shows the selected

Table 2 Hyper-parameters of the build CNN model and preferred weights in this study

R1	R2	R3	R4	R5	R6	R7	R8
CNN	224*224	RMSprop	32	10-fold	3e-4	BCE	50

R1—model, R2—image size, R3—optimizers, R4—mini batch size, R5—cross validation, R6—initial learning rate, R7—loss function, R8—Epoch, *BCE*—binary cross-entropy

hyperparameters. There are five convolution layers in this proposed CNN model, which take as its input a retinal fundus image tensor of 244×244.

Subsequently, the first convolution layer uses $5 \times 5 \times 3$ kernel filters with stride 1×1, and a total of 64 such filters are employed. The next layer, which receives the output from the first layer, is a max-pooling layer with 2×2 stride, reducing the input to half of its size 112×112. For all layers, the output from the pooling layer passes through the ReLU activation feature. The nonlinear output obtained is now fed into the next convolution layer with $5 \times 5 \times 64$ with 128 filters, and the stride value is the same 1×1. The obtained output pass through a max-pooling layer with the same 2×2 strides, which again reduced the input to half of its size 56×56. After the output pass through ReLU activation, it is fed into the third convolution layer with 256 filters and the kernel size $5 \times 5 \times 128$ with 1×1 stride. The output is passed to a max-pooling layer, which results in a tensor of shape 28×28. Again the output pass through ReLU activation, fed into the fourth convolution layer with 512 filters and kernel size $5 \times 5 \times 256$ and with the same stride 1×1. The output from the fourth convolution is max-pooled to a size of 14×14. After ReLU activated and it is pass to a fifth convolution layer with 512 filters and $14 \times 14 \times 512$ kernel size to accommodate the output of all the filters from previously configured layers, and max-pooling of output from that layer with a stride of size 2×2 produces an output of size 14×14. Now the resulting tensor has the shape $7 \times 7 \times 512$. The obtained tensor is flattened with 25,088 neurons. The weighed values that emerge as neurons demonstrate the proximity to the symptoms of COVID-19. The dropout layer is applied here to drop values to handle network overfitting. In our work, we used a dropout rate of 0.5 during training. The fully connected layer converts the tensor with 25,088 neurons to 64 neurons and adds ReLU activation to the output. A tensor with 64 neurons is the product of the fully connected layers; these 64 neurons are translated into neuron counts equal to the number of categories to which the retinal image belongs, healthy, diabetic retinopathy, glaucoma, and diabetic macular edema.

3.6 Evaluation Criteria

Different metrics have been used to evaluate the efficiency of the highest performing DL model. To calculate the true or false classification of the DED diagnosed in the fundus images evaluated as follows. Initially, the cross-validation estimator [45] is adopted and plotted in a confusion matrix as shown in Fig. 5. The confusion matrix has the following four predicted outcomes. True Positive (TP) has been identified with the right diagnosis and a variety of abnormalities. True Negative (TN) is an erroneously calculated number of periodic instances. False positives (FP) are a set of periodic instances. The following performance metrics are used to calculate the values of possible outcomes in the confusion matrix.

Accuracy Accuracy is an essential metric for the evaluation of the results of DL classifiers. It is a summary of the true positive and true negatives divided by the confusion of the matrix components' total values. The most accurate model is an excellent one, but it is imperative to ensure that symmetric sets of data with almost equal false positive values and false negative values. Thus, the elements of the confusion matrix mentioned above will be calculated to evaluate the effectiveness of our proposed classification model for the DED dataset.

$$\text{Accuracy (\%)} = \frac{TP + TN}{TP + FN + TN + FP} 100\%. \tag{22}$$

Sensitivity (Recall): sensitivity is measured as the number of accurate positive predictions divided by the sum of positive. The best sensitivity is 1.0, whereas the worst is 0.0. We calculate sensitivity using following equation;

$$\text{Sensitivity} = \frac{TP}{TP + FN} \tag{23}$$

Specificity: specificity is measured as the number of correct negative predictions divided by the sum of negatives. The

	P Positive	P Negative	Total
A Positive	*TP*	*FN*	*TP + FN*
A Negative	*FP*	*TN*	*FP + TN*
Total	*Se*	*Sp*	

Fig. 5 Confusion matrix: a positive = actual positive, a negative = actual negative, P positive = predicted positive, P negative = predicted negative, *TP* true positive, *FN* false negative, *FP* false positive, *TN* true negative, Se = TP + FP, Sp = FN + TN

Table 3 Average performance of the models on original images

DED	Model	Accuracy (%)	Sensitivity (%)	Specificity (%)	Precision (%)
Normal /mild DR	Xception	60.87	67	58	43
	VGG16	**80.43**	**76.92**	**85**	**74**
	DenseNet121	56.67	71.43	53.85	96
Normal /mild DME	Xception	62.07	65	60	52
	VGG16	**85.79**	**90**	**81**	**78**
	DenseNet121	51.72	100	51	28
Normal /mild GL	Xception	63.41	85.71	58.82	95
	VGG16	**87.80**	**94.12**	**83.33**	**95**
	DenseNet121	80.49	77	84	76

Bold values indicate VGG16 classification surpassed the other two fully-trained deep learning models

DED diabetic eye disease, *DR* diabetic retinopathy, *DME* diabetic macular edema, *Gl* Glaucoma

best specificity is 1.0, whereas the worst is 0.0. We calculate sensitivity using the following equation;

$$Specificity = \frac{TN}{TN + FP} \tag{24}$$

4 Experiments and Results

All the experiments are implemented using MatLab, Python, Keras library1, with TensorFlow2 as a back-end and Python 3.8 programming language in jupyter notebook with a processor 2.3 GHz Intel Core i9 and RAM of 16 GB 2400 MHz DDR4 with Intel UHD Graphics 630 1536 MB. The training/testing data split was set at 80/20. The segregated generic selection was conducted to ensure an approximately equal distribution of the class. Mini-batch size was set to 32, and the cross-entropy loss function was chosen due to its suitability for binary classification tasks. The Optimiser was set as default (Adam) and RMSprop for build CNN. The standard performance evaluation metric accuracy, sensitivity, and specificity of the test dataset were used to validate results.

We compared and analysed performance accuracy among three distinct pre-trained deep learning models with the new build CNN model in this study. The three pre-trained models, namely; *Xception*, *VGG16*, and *DenseNet21* and five-layered convolutional model have been evaluated in terms of test data set accuracy (Table 1). The pre-trained models adopted for this research were trained and tested by large-scale ImageNet data, covering a wide range of categories such as cars, animals, flowers, etc. Models acquire excellent performance image classification for objects while demonstrating a limitation in their application to narrow product areas, such as medical lesion (DED) detection. The prognosis of pathological indications in the retinal fundus images depends upon various complex characteristics and lesion localization in the retinal fundus image. There is a new representation of the input image in each CNN layer by progressive extraction of the most distinctive features. For instance, the first layer is capable of learning edges, while the last layer can identify a lesion as a DED classification feature. As a result, the following scenarios were considered in the experiments: Region of Interest such as blood vessels, macular regions, and the optic disc has been detected, localized, and segmented.

We have employed a combination of multiple traditional image segmentation algorithms for each phase in the proposed system. All of these algorithms provided effective results in the segmentation of the region of interest. We performed a series of procedures to build a high-performance system, such as image enhancement, blood vessel segmentation, identification and then extraction of optic discs, extraction of macular region, blood vessels removal, elimination of optic discs, extraction of features, and classification of features. After segmentation, the size of the images has been optimized to a suitable size following the input specifications of each network. To minimise the risk of model overfitting, the imbalance dataset was augmented using real-time augmentation *ImageDataGenerator* class from Keras. The fine-tuning was used for pre-trained models after eliminating and re-training *n* layers (*n* was CNN layer-dependent). The final output acquired for each model was used for comparison in terms of percentage accuracy are represented in Tables 3 and 4. VGG16 classification surpassed the other two fully-trained deep learning models Xception and DenseNet121. Similarly, among all pre-trained models, new build CNN model using preprocessed retinal images performed well. Tables 5 and 6 compare the accuracy of results. Build CNN accuracy surpasses all the model used for classifications.

To detect retinal anomalies, we developed more general screening classification models. The confusion matrix and

Table 4 Average performance of the VGG16 model on pre-processed images

DED	Model	Accuracy (%)	Sensitivity (%)	Specificity (%)	Precision (%)
Normal /mild DR	VGG16	83.43	86	85.71	78
Normal /mild DME	VGG16	89.13	85	95	96
Normal /mild GL	VGG16	88	95	90	90

DED diabetic eye disease, *DR* diabetic retinopathy, *DME* diabetic macular edema, *Gl* Glaucoma

Table 5 Average performance of the new proposed model on original images

DED	Model	Accuracy (%)	Sensitivity (%)	Specificity (%)	Precision (%)
Normal /mild DR	CNN	63.33	53.33	73.33	61
Normal /mild DME	CNN	82.86	83.33	82.35	82
Normal /mild GL	CNN	96.77	100	93.75	100

DED diabetic eye disease, *DR* diabetic retinopathy, *DME* diabetic macular edema, *Gl* Glaucoma

Table 6 Average performance of the new proposed model on pre-processed Images

DED	Model	Accuracy (%)	Sensitivity (%)	Specificity (%)	Precision (%)
Normal /mild DR	CNN	93.33	100	86.67	100
Normal /mild DME	CNN	91.43	94.44	88.24	94
Normal /mild GL	CNN	100	100	100	100

DED diabetic eye disease, *DR* diabetic retinopathy, *DME* diabetic macular edema, *Gl* Glaucoma

Fig. 6 VGG16 model performance in diabetic retinopathy

(a) Confusion Matrix   **(b)** ROC-AUC

(a) Confusion Matrix

(b) ROC-AUC

Fig. 7 VGG16 model performance in diabetic macular edema
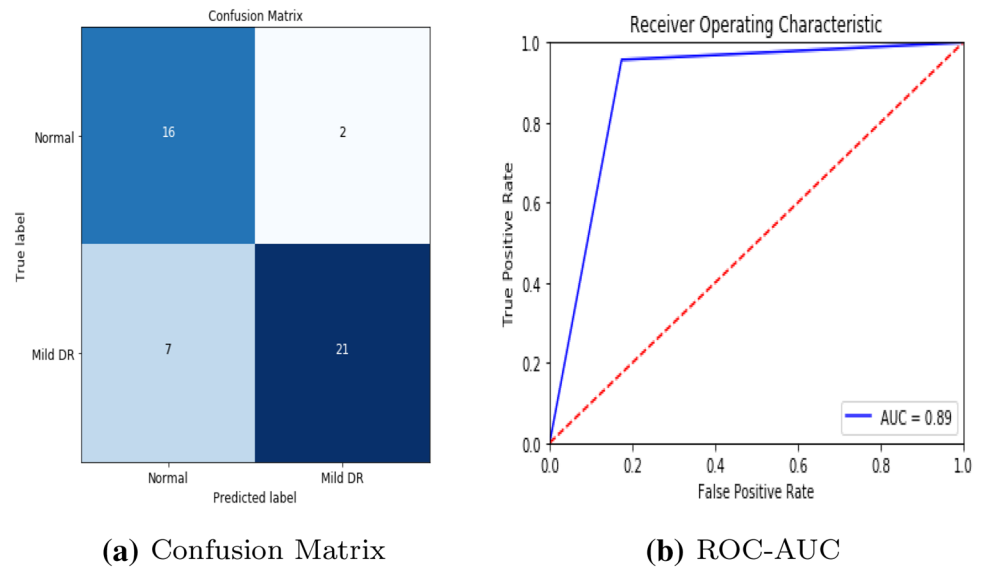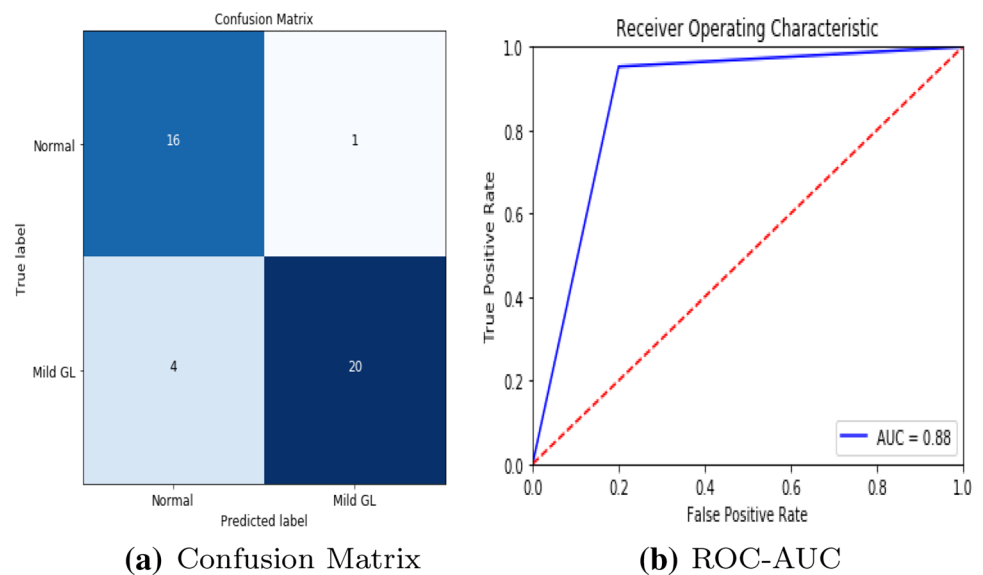

(a) Confusion Matrix

(b) ROC-AUC
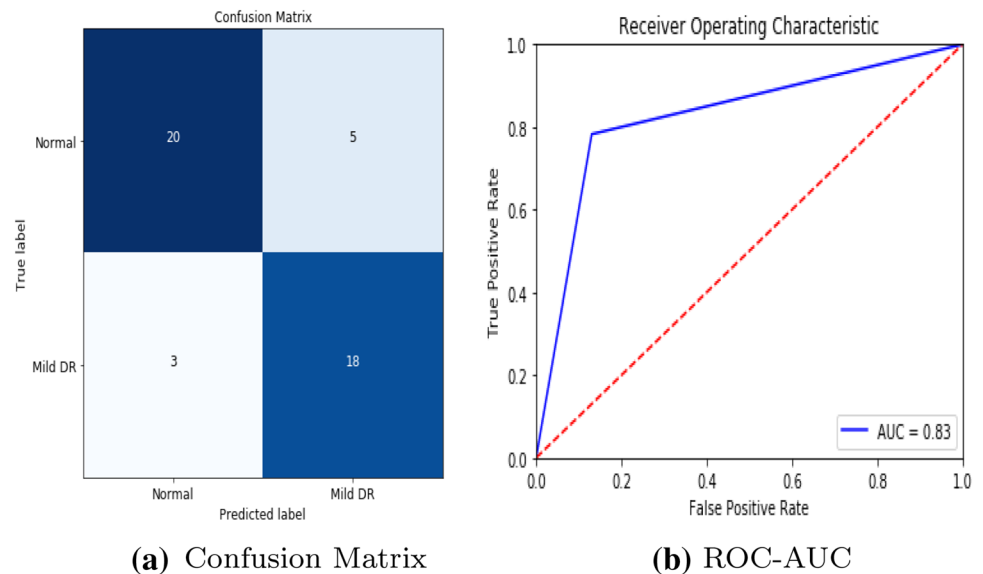
Fig. 8 VGG16 model performance in Glaucoma

ROC curves of each pre-trained deep learning model and a build CNN model for binary classification of healthy and other DED disease status are shown in Figs. 6, 7, 8, 9, 10 and 11.

5 Discussions

This research is a study of binary classification deep learning algorithms to identify three mild diabetic eye diseases automatically. This research has shown that the complexity of the deep learning algorithms arises from the quality and quantity of data (fundus images), not from the algorithm. In this research, we used publicly available annotated data (fundus images). For a computer-aided clinical application, more robust, practical, and realistic results can be obtained using labeled hospital fundus images. Indeed, this paper recommends that the automatic classifier strive to classify against the binary classification of at least average, DR, DME, and GL due to each disease's significance. These three diseases are the major retinal diseases caused by diabetes; unless an initial evaluation is conducted, these diseases always cause severe damage to the visual acuity, and it is irreversible [6, 38].

Growing life expectancy, hustling lifestyles, and other factors suggest that the number of people with diabetes is expected to increase [38]. For example, many patients with DED in China often overlook their situation and lack timely treatments leading to serious state development of DR [55]. Early intervention of abnormal signs prevents further

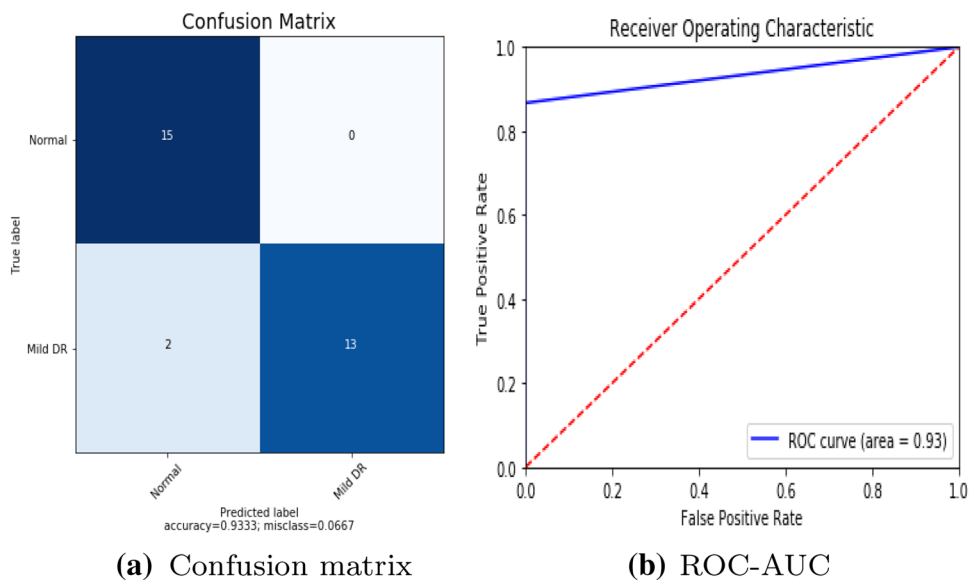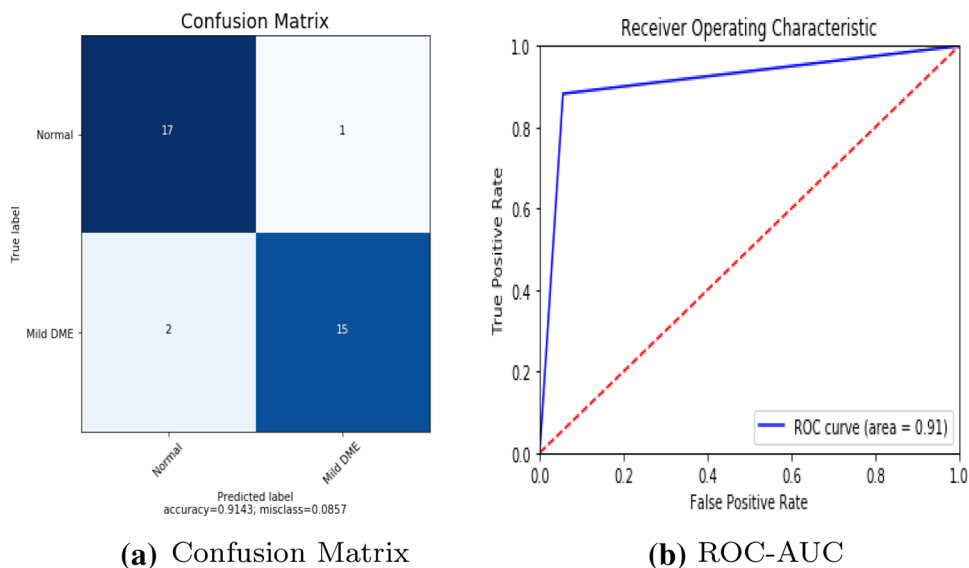**Fig. 9** Build CNN model performance in diabetic retinopathy

(a) Confusion matrix

(b) ROC-AUC

Fig. 10 Build CNN model performance in diabetic macular edema

(a) Confusion Matrix

(b) ROC-AUC

deterioration of the condition and its effect on the impacted individuals and related medical costs. Therefore, the DED identification system enables either completely automate the eye-screening process or semi-automate eye-screening system. The first method requires a reasonable degree of accuracy, which is similar to that of the retinal experts. As per the British Diabetic Association (BDA) guidelines, a minimum level of 95% specificity and 80% sensitivity for sight-threatening DR detection must be obtained applied method [3]. Second option allows to downsize the large-scale mass-screening outputs to the potential DED cases, followed by human examination. Both scenarios significantly reduce the burden on skilled ophthalmologists and specialised facilities, making the process accessible to wider population, especially in low-resource settings.

After all, the application of deep learning to the clinical practice still has many challenges. An earlier paper dealt with ethical and political concerns in terms of database creation [1]. For this purpose, it has been difficult to obtain large-scale data on many diabetic eye diseases. Another challenge is that mild (early) classification problems consist of real clinical problems. Binary classification for diabetic eye disease prediction was the subject of previous studies. Even if Google has built a deep learning model that works better than ophthalmologists, their 'Inception-v3' model was optimised for binary classification for DR identification based on the GoogLeNet structure Gulshan et al. [17]. This framework was evaluated after adding a wide image database gathered for only healthy and non-healthy DR screening from diabetes patients. For binary

Fig. 11 Build CNN model
performance in Glaucoma

(a) Confusion Matrix   **(b)** ROC-AUC
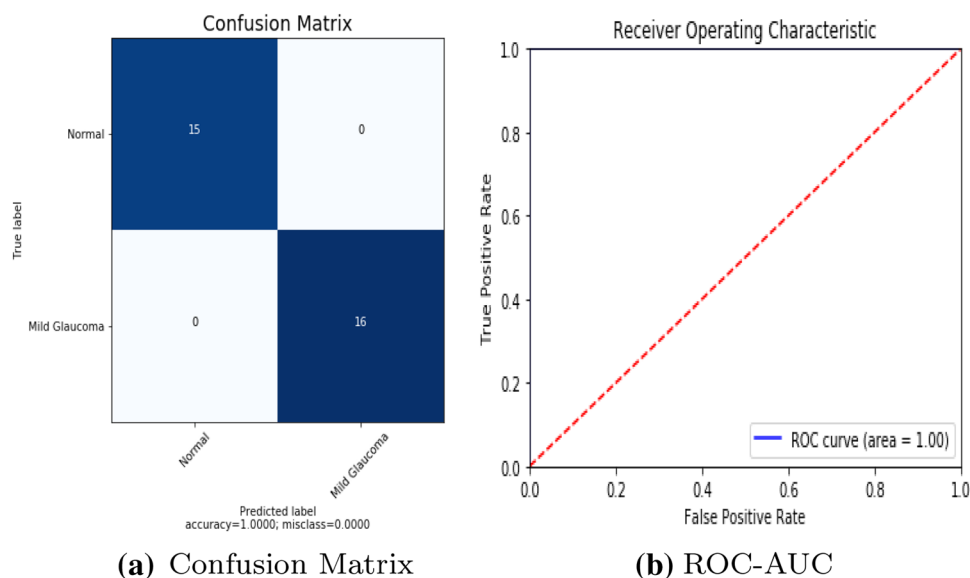
disease classification, Gulshan et al. stated a 93-96 percent recall but noted that recall is not enhanced while practicing with 60,000 image samples contrast with 120,000 image samples employing a private data.

Visual representations of the features acquired by CNNs demonstrate that the patterns being used for classification are a part of the image fully visible to the observer [53]. Moderate and severe class of the diabetic retinal images include macroscopic features on a scale designed for classification by current CNN architectures, such as those accessible from the ImageNet visual database. On the other hand, less than 1 % of the overall pixel volume, a degree of slight that is often difficult for human interpreters to identify, is the characteristics that differentiate mild from the normal disease.

This research indicates that mild-class DED classification should be established through further studies on automatic diagnosis using retinal fundus imaging. The first part of the experiment includes traditional image processing for enhancing mild DED features. Various conventional techniques for image processing have been implemented to extract DED lesions. Pre-trained CNN models using transfer learning provides excellent performance with object-oriented images such as flowers, cars, and animals but not efficient with lesion based medical images. So in this research we aim to objectify mild DED lesion via segmenting region of interest and transfer it to transfer learning and build CNN for further feature extraction. Following the elimination of the top layer (existing approach). A detailed review of 3 CNN architectures (including state-of-the-art architectures) was conducted. Secondly, the n layers were 'unfrozen' and then re-trained to respond effectively to the details of the case-study of the application (proposed approach). Messidor, Messidor-2, retinal dataset were provided for system

training. Two training sets were prepared with available dataset (i) before, and (ii) after preprocessing to measure potential accuracy improvement for Normal/Mild DED image classification.

As Mild DED tends to be incredibly difficult to discern from a normal retina due to few subtle indication of the impairment, an increase in the quality data was supposed to improve the visibility of pathological features. The top 1 CNN architectures with the top layer removed and re-trained were *VGG16*, yielding the accuracy of 83.43%, 89.13%, 88% for each (Table. 4). The lowest performance was obtained by *Xception* and *DenseNet21* respectively. The impact of fine tuning varied across the models. The observed improvement in accuracy was only minor, indicating the relative appropriateness of default pre-trained networks for DED classification tasks. In other terms, the CNN networks were able to identify Mild DED from a healthy retina despite having been trained on different images from the ImageNet repository. If no improvement in accuracy is obtained, the unfreezing is not advised result in unnecessary computational costs and time accrued. For build CNN model yielded the accuracy of 93.33%, 91.43%, 100% respectively.

For comparing the performance of the employed models, the 2 scenarios were considered, (1) before image preprocessing, and (2) after image preprocessing. In before preprocessing scenario, we trained our models with a raw dataset with data augmentation (geometric transformation) were applied to Messidor, Messidor-2, DRISTI-GS dataset, to avoid overfitting. In after image preprocessing scenario, dataset were preprocessed using various traditional image processing techniques which has increased the classification performance to 100% (the maximum accuracy achieved for Gl).

After evaluation of our high performed approach on Mild DR, Mild DME and Mild GL detection task, the maximum sensitivity of 100%, 94.44%, 100% and the maximum specificity of 86.67%, 88.24%, 100% were obtained. Thus, the early DED detection proved sufficient given the BDA standards, but still falling 9% and 6% short in terms of its specificity.

5.1 Approach Limitations

Several deficiencies of the research have been established. First, Data set acquired for this experiment were obtained from publicly available which limits number of high quality mild DED images, only limited-to-moderate data set sizes were employed in the research. The approach also emphasises the value of an effective annotation process as having a direct effect on the output of the classifier. The Messidor, Messidor-2, retinal dataset has been validated and marked by professional ophthalmologists. Transfer Learning is used as a compensation procedure. Pre-trained CNN models in the wide-scale ImageNet database have been adopted in this study. To increase the size of the training sample set and to ease the data imbalance problem data were rotate, flipped, mirrored, etc. Second, the default model parameters were adopted for classification task (i.e. dropout, batch size, loss function, optimizer etc.). Finally, the 'black-box' nature of Deep Learning-based solution is often criticized, causing resistance in the broader approach adopted by practitioners. However, with a build CNN using binary classifiers, we achieved state-of-the-art accuracy, the model performance degrades with the use of transfer learning. However, it is striving to ensure that more data might be more robust. Previous field research has confirmed that CNN's ability to accommodate differences in size is limited and some have indicated that more data can not complement this inherent weakness in the case of retinal images [17].

6 Conclusion

This paper proposes an approach that focuses specifically on the identification of mild DED among Normal instances as not adequately discussed in previous literature. According to the analytical aspect of Deep Learning, a variety of performance optimization techniques have been employed (1) Image enhancement, (2) feature enhancement, (3) data balance, and (4) fine-tune. The additional advantage of Deep Learning involves automatic recognition capabilities that are most selective between categories. Such an approach makes it possible to avoid technological constraints with the analytical, and sometimes subjective, approach of manual extraction of features. In addition, the analysis used composite data sets from various sources to determine the robustness

of the system and its capacity to respond to real-world scenarios. The developed system enables the standardisation of labour-intensive eye-screening processes and satisfies as an auxiliary diagnosing reference, while avoiding human subjectivity.

References

1. Abramoff MD, Niemeijer M, Russell SR (2010) Automated detection of diabetic retinopathy: barriers to translation into clinical practice. Expert Rev Med Devices 7(2):287–296
2. Acharya UR, Lim CM, Ng EYK, Chee C, Tamura T (2009) Computer-based detection of diabetes retinopathy stages using digital fundus images. Proc Inst Mech Eng Part H J Eng Med 223(5):545–553
3. Association BD et al (1997) Retinal photography screening for diabetic eye disease. BDA, London
4. Bargshady G, Zhou X, Deo R, Soar J, Whittaker F, Wang H (2020) Enhanced deep learning algorithm development to detect pain intensity from facial expression images. Expert Syst Appl 149:113305. https://doi.org/10.1016/j.eswa.2020.113305
5. Bezdek J, Pal MR, Keller J, Krisnapuram R (1999) Fuzzy models and algorithms for pattern recognition and image processing, 4th edn. Springer
6. Bond P. South africas vulnerability to the world capitalist crisis: how it worsened and how it might be reversed. OF THE NATION, p 609
7. Briot JP, Hadjeres G, Pachet FD (2017) Deep learning techniques for music generation—a survey. arXiv preprint arXiv:1709.01620
8. Chaudhuri S, Chatterjee S, Katz N, Nelson M, Goldbaum M (1989) Detection of blood vessels in retinal images using

two-dimensional matched filters. IEEE Trans Med Imaging 8(3):263–269

9. Chollet, F. (2017) Xception: Deep learning with depthwise separable convolutions. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 1251–1258

10. Coye T (2015) A novel retinal blood vessel segmentation algorithm for fundus images. MATLAB Central File Exchang

11. Du J, Michalska S, Subramani S, Wang H, Zhang Y (2019) Neural attention with character embeddings for hay fever detection from twitter. Health Inf Sci Syst 7(1):1–7

12. Du J, Zheng L, He J, Rong J, Wang H, Zhang Y (2020) An interactive network for end-to-end review helpfulness modeling. Data Sci Eng 5(3):261–279

13. Gao W, Peng M, Wang H, Zhang Y, Xie Q, Tian G (2018) Incorporating word embeddings into topic modeling of short text. Knowledge and Information Systems, pp 1–23

14. Gardner GG, Keating D, Williamson TH, Elliott AT (1996) Automatic detection of diabetic retinopathy using an artificial neural network: a screening tool. Br J Ophthalmol 80(11):940–944

15. Gargeya R, Leng T (2017) Automated identification of diabetic retinopathy using deep learning. Ophthalmology 124(7):962–969

16. Gharaibeh NY (2017) A novel approach for detection of microaneurysms in diabetic retinopathy disease from retinal fundus images. Comput Inf Sci 10(1):1–15

17. Gulshan V, Peng L, Coram M, Stumpe MC, Wu D, Narayanaswamy A, Venugopalan S, Widner K, Madams T, Cuadros J et al (2016) Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. Jama 316(22):2402–2410

18. He J, Rong J, Sun L, Wang H, Zhang Y, Ma J (2020) A framework for cardiac arrhythmia detection from iot-based ecgs. World Wide Web 23:2835–2850. https://doi.org/10.1007/s11280-019-00776-9

19. Hu H, Li J, Wang H, Daggard G (2006) Combined gene selection methods for microarray data analysis. Knowledge-based intelligent information and engineering systems. Springer, Berlin, Heidelberg, pp 976–983

20. Huang G, Liu Z, Weinberger KQ (2017) Densely Connected Convolutional Networks. In: 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp 2261–2269

21. Hutchinson A, McIntosh A, Peters J, Okeeffe C, Khunti K, Baker R, Booth A (2000) Effectiveness of screening and monitoring tests for diabetic retinopathy—a systematic review. Diabet Med 17(7):495–506

22. Jiang H, Zhou R, Zhang L, Wang H, Zhang Y (2019) Sentence level topic models for associated topics extraction. World Wide Web 22:2545–2560. https://doi.org/10.1007/s11280-018-0639-1

23. Juneja M, Singh S, Agarwal N, Bali S, Gupta S, Thakur N, Jindal P (2019) Automated detection of glaucoma using deep learning convolution network (g-net). Multimed Tools Appl pp 1–23

24. Karegowda AG, Nasiha A, Jayaram M, Manjunath A (2011) Exudates detection in retinal images using back propagation neural network. Int J Comput Appl 25(3):25–31

25. Kaur M, Kaur M (2015) A hybrid approach for automatic exudates detection in eye fundus image. Int J 5(6):411–417

26. Khalil F, Wang H, Li J (2007) Integrating markov model with clustering for predicting web page accesses. In: The 13th Australasian world wide web conference, pp 63–74

27. Lam C, Yi D, Guo M, Lindsey T (2018) Automated detection of diabetic retinopathy using deep learning. AMIA Summit Translat Sci Proc 2018:147

28. Li H, Wang Y, Wang H, Zhou B (2017) Multi-window based ensemble learning for classification of imbalanced streaming data. World Wide Web 20:1–19. https://doi.org/10.1007/s11280-017-0449-x

29. Litjens G, Kooi T, Bejnordi BE, Setio AAA, Ciompi F, Ghafoorian M, Van Der Laak JA, Van Ginneken B, Sánchez CI (2017) A survey on deep learning in medical image analysis. Med Image Anal 42:60–88

30. Mnih V, Kavukcuoglu K, Silver D, Rusu AA, Veness J, Bellemare MG, Graves A, Riedmiller M, Fidjeland AK, Ostrovski G et al (2015) Human-level control through deep reinforcement learning. Nature 518(7540):529–533

31. Mookiah MRK, Acharya UR, Chua CK, Lim CM, Ng E, Laude A (2013) Computer-aided diagnosis of diabetic retinopathy: a review. Comput Biol Med 43(12):2136–2155

32. Noronha K, Nayak J, Bhat SN (2006) Enhancement of retinal fundus Image to highlight the features for detection of abnormal eyes. TENCON 2006-2006 IEEE Region 10 Conference, pp 1–4

33. Osareh A, Mirmehdi M, Thomas B, Markham R (2003) Automated identification of diabetic retinal exudates in digital colour images. Br J Ophthalmol 87(10):1220–1223

34. Pan SJ, Yang Q (2009) A survey on transfer learning. IEEE Trans Knowl Data Eng 22(10):1345–1359

35. Pandey D, Yin X, Wang H, Zhang Y (2016) Accurate vessel segmentation using maximum entropy incorporating line detection and phase-preserving denoising. Computer Vision and Image Understanding. 155. https://doi.org/10.1016/j.cviu.2016.12.005

36. Peng M, Zeng G, Sun Z, Huang J, Wang H, Tian G (2018) Personalized app recommendation based on app permissions. World Wide Web 21:1–16. https://doi.org/10.1007/s11280-017-0456-y

37. Peng M, Zhu J, Wang H, Li X, Zhang Y, Zhang X, Tian G (2018) Mining event-oriented topics in microblog stream with unsupervised multi-view hierarchical embedding. ACM Trans Knowl Discov Data 12:1–26. https://doi.org/10.1145/3173044

38. Pratt H, Coenen F, Broadbent DM, Harding SP, Zheng Y (2016) Convolutional neural networks for diabetic retinopathy. Procedia Comput Sci 90:200–205

39. Sarki R, Ahmed K, Wang H, Zhang Y (2020) Automated detection of mild and multi-class diabetic eye diseases using deep learning. Health Inf Sci Syst 8(1):1–9

40. Sarki R, Ahmed K, Wang H, Zhang Y (2020) Automatic detection of diabetic eye disease through deep learning using fundus images: a survey. IEEE Access 8:151133–151149. https://doi.org/10.1109/ACCESS.2020.3015258

41. Shih FY (2010) Image processing and pattern recognition: fundamentals and techniques. Wiley

42. Silver D, Huang A, Maddison CJ, Guez A, Sifre L, Van Den Driessche G, Schrittwieser J, Antonoglou I, Panneershelvam V, Lanctot M (2016) Mastering the game of go with deep neural networks and tree search. Nature 529(7587):484–489

43. Simonyan K, Zisserman A (2015) Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv: 1409.1556

44. Sivaswamy J, Krishnadas SR, Datt Joshi G, Jain M, Syed Tabish AU (2014) Drishti-GS: Retinal image dataset for optic nerve head(ONH) segmentation. In: 11th International Symposium on Biomedical Imaging (ISBI), pp. 53–56. https://doi.org/10.1109/ISBI.2014.6867807

45. Sokolova M, Lapalme G (2009) A systematic analysis of performance measures for classification tasks. Inf Process Manag 45(4):427–437

46. Solomon C, Breckon T (2011) Fundamentals of digital image processing: a practical approach with examples in Matlab. Wiley

47. Sopharak A, Uyyanonvara B (2007) Automatic exudates detection from diabetic retinopathy retinal image using fuzzy c-means and morphological methods. In: Proceedings of the third IASTED international conference advances in computer science and technology, pp 359–364

48. Supriya S, Siuly S, Wang H, Zhang Y (2020) Automated epilepsy detection techniques from electroencephalogram signals: a review study. Health Inf Sci Syst 8(1):1–15

49. Tan WR, Chan CS, Aguirre HE, Artgan TK (2017) Artwork synthesis with conditional categorical gans. IEEE international conference on image processing (ICIP), pp 3760–3764

50. Taylor R, Batey D. Handbook of retinal screening in diabetes. Wiley (2006)

51. Vallabha D, Dorairaj R, Namuduri K, Thompson H (2004) Automated detection and classification of vascular abnormalities in diabetic retinopathy. In: 28th asilomar conference on signals, systems and computers,vol. 2, pp 1625–1629

52. Yin J, Tang M, Cao J, Wang H, You M, Lin Y (2020) Adaptive online learning for vulnerability exploitation time prediction. In: International Conference on Web Information Systems Engineering, pp 252–266.Springer

53. Yosinski J, Clune J, Nguyen A, Fuchs T, Lipson H (2015) Understanding neural networks through deep visualization. arXiv preprint arXiv:1506.06579

54. Youssif AA, Ghalwash AZ, Ghoneim AS, et al (2006) Comparative study of contrast enhancement and illumination equalization methods for retinal vasculature segmentation. Cairo international biomedical engineering conference, pp 1–5

55. Zhang W, Zhong J, Yang S, Gao Z, Hu J, Chen Y, Yi Z (2019) Automated identification and grading system of diabetic retinopathy using deep neural networks. Knowl-Based Syst 175:12–25

56. Zuiderveld K (1994) Contrast limited adaptive histogram equalization. Graphics gems IV, pp 474–485. Academic Press Professional, Inc